Europäisches Patentamt
European Patent Office
Office européen des brevets

(19)

(11) Publication number: **0 424 162 A2**

EUROPEAN PATENT APPLICATION

(12)

(21) Application number: 90311472.6

(51) Int. Cl.[5]: H04B 1/66

(22) Date of filing: 18.10.90

(30) Priority: 18.10.89 JP 271011/89

(43) Date of publication of application: 24.04.91 Bulletin 91/17

(84) Designated Contracting States: DE FR GB

(71) Applicant: VICTOR COMPANY OF JAPAN, LTD.
12, Moriya-Cho 3-Chome, Kanagawa-Ku
Yokohama-Shi, Kanagawa-Ken(JP)

(72) Inventor: Fuchigami, Tokuhiko, A-203, Victor Tachibanadai
Dome, 2-6-1, Tachibanadai
Midori-Ku, Yokohama-Shi,
Kanagawa-Ken(JP)
Inventor: Konishi, Masaya
B-101, Nippon Victor Nobi-Ryo, 2164, Nagasawa
Yokosuka-Shi, Kanagawa-Ken(JP)
Inventor: Yasura, Sadahiro
2-18-12, Omori-Nishi
Ota-Ku, Tokyo-To(JP)
Inventor: Yamada, Yasuhiro
304 Tahachi Bldg., 27, Nagasawa
Yokosuka-shi, Kanagawa-Ken(JP)

(74) Representative: Crawford, Andrew Birkby et al
A.A. THORNTON & CO. Northumberland House 303-306 High Holborn
London WC1V 7LE(GB)

(54) Method of coding an audio signal by using an adaptive orthogonal transformation.

(57) A method of coding an audio signal by using an adaptive orthgonal transformation comprises a step (ST1, ST2) of calculating the sum total of a power of an input audio signal in a time duration of an orthogonal transform length or over the length; a step (ST3) of adaptively controlling the minimum audible level as a threshold level corresponding to the power level on a frequency axis after the orthogonal transformation or on a time axis before them orthogonal transformation; a step (ST4) of excepting samples having a power level under the threshold level from orthogonally transformed samples; and a step (ST5) of quantizing remaining samples. By the method, the threshold level of the minimum audible level is adaptively controlled corresponding to an observed level of an input audio signal. The method has an effect that even if the input audio signal has various power levels, it is possible to usually set the adaptively audible value (the threshold level), thereby effectively reducing a coding amount.

EP 0 424 162 A2

METHOD OF CODING AN AUDIO SIGNAL BY USING AN ADAPTIVE ORTHOGONAL TRANSFORMATION

BACKGROUND OF THE INVENTION

The present invention relates to a method of coding an audio signal by using orthogonal transform by using an adaptive transformation such as a discrete cosine tansformation (DCT), a discrete Fourier transformation (DFT) and the like, and more particularly to a method in which a minimum audible acoustic level is set to a predetermined value adaptive to a power level for reducing a coding rate.

When an audio signal is compressed and coded, a conventional system excepts "an unnecessary component" of the signal component as being inaudible for effectively reducing a code rate.

A person having normal hearing ability can generally hear sounds within the area enclosed by two lines representing the maximum and minimum audible levels, respectively. This audible area is illustrated in FIG. 1 as an area with hatching. Accordingly, signals representing sounds under the minimum audible level can be ignored.

A signal on the time axis is transformed on a frequency axis by the orthogonal transformation and this resultant coefficient may distribute from a full scale value (if 16 bit data, it is 2^{15}) to an operational accuracy value. In the conventional method, the full scale value is set to the maximum audible level (120 dB Sound Pressure Level -S.P.L.-) and the minimum audible level (about 0 dB S.P.L. at 1kHz) is determined according to the maximum level, as shown in FIG. 1, and samples (coefficient) under the minimum audible level are omitted. However, it is very rare for actual music signals to hit an upper value of the full scale.

Accordingly, when a signal source having an average of a signal level at about 60 dB is set to 80 dB S.P.L. of an average at reproduction, as shown in FIG. 2(a), an oblique lined portion omitted in coding as inaudible sample, is as shown in FIG. 2(b) at an actual reproduction, so that a part of the actual audible portion is lost.

In this case, the minimum audible level (hereinafter, the level is called a threshold level) needs to be set to be lowered by 20 dB. In contrast, when the signal of 80 dB of an average level is reproduced by 80 dB S.P.L. under the threshold level, is set before, as shown in FIGS. 3(a) and 3(b), there remains signals which cannot be actually audible.

As described above, a conventional method has problems that too many signals are lost when the level of reproduction is set higher, and in contrast, unnecessary signals are left when the level of reproduction is set lower.

SUMMARY OF THE INVENTION

An object of the present invention is to provide a method of coding an audio signal, capable of effectively coding the audio signal by adaptively or variably controlling a threshold level as a set value of the minimum audible level corresponding to an actual signal level previously observed.

In order to achieve the above object, the method of the present invention comprises steps of calculating the sum total of a power of an input audio signal in a time duration of an orthogonal transform length or over the length, adaptively controlling the minimum audible level as a threshold level corresponding to the power level on a frequency axis after the orthogonal transformation, excepting samples having a power level under the threshold level from orthogonally transformed samples, and quantizing the remaining samples.

By this invention, the threshold level of the minimum audible level is adaptively controlled corresponding to an observed level of an input audio signal.

As is apparent from the above-mentioned, the present invetion has the effect that even if the input audio signal has various power levels, it is possible to usually set the adaptive audible value (the threshold level), thereby effectively reducing the coding amount.

Furthermore, it is a feature that the coding amount is smoothed at the same time as coding by means of keeping a substantially fixed dynamic range of each interval (the orthogonal transformation length).

BRIEF DESCRIPTION OF THE DRAWINGS

In the accompanying drawings:

FIG. 1 is a characteristic diagram showing the typical audible range of a person having normal hearing ability;

FIGS. 2(a) and 2(b) are diagrams showing how a signal being inaudible at coding results in a signal being audible at reproduction;

FIGS. 3(a) and 3(b) are diagrams showing how a signal being audible at coding results in a signal being inaudible at reproduction;

FIG. 4 is a block diagram showing a coding unit for utilizing a coding method of the present invention;

FIG. 5 is a flow chart showing a method of coding an audio signal by using an adaptive orthogonal transformation according to a example of the present invention;

FIG. 6 is a characteristic diagram showing the concept for processing according to the present invetion; and

FIG. 7 is a polygonal line graph showing a threshold power level as a function of frequency.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

There will now be described in detail a method of coding an audible signal by using an adaptive orthogonal transformation according to an embodiment of the present invention with reference to the attached drawings.

FIG. 1 is a block diagram showing a coding unit for utilizing a coding method of this invention, and FIG. 2 is a flow chart showing an operational procedure of the coding method of this invention.

[Configuration of the Coding Unit]

The coding unit for utilizing this invention comprises, as shown in FIG. 4, a window circuit 1 for dividing an input audio signal into successive blocks having a predetermined length, an orthogonal transformation circuit 2 such as the discrete cosine transformation (DCT) or the discrete Fourier transformation (DFT), a power calculation circuit 3 for calculating a power level of the input audio signal, an acoustic characteristic data memory 4 for storing acoustic characteristic data, a cut-off circuit 5 for cutting off a signal component under a threshold level, a quantization circuit 6 for quantizing a remaining component without the cut-off component, and a coding circuit 7 for coding the remaining component.

[Procedure of the Coding Method]

The coding method is performed by following steps as shown in FIG. 5.

In the window circuit 1, the input audio signal is divided into successive blocks having a predetermined length for orthogonal transformation and supplied to both the orthogonal transformation circuit 2 (DCT or DFT) and the power calculation circuit 3 each connected the window ciruclt 1 (refer to step ST1).

Each block so divided is orthogonally transformed in the transformation circuit 2.and supplied to the cut-off circuit 5. The power calculation circuit 3 calculates the power level of the input audio signal for a predetermined period equal to one block length (orthogonal transformation length) or more (refer to step ST2).

The power level calculated above is supplied to the cut-off circuit 5 with the data output from the acoustic data memory 4.

The minimum audible level (a threshold level) is adaptively controlled on a frequency axis after the orthogonal transformation (refer to step ST3). As will be described later, the cut-off circuit 5 controls the threshold levels A and B corresponding to the input power level A and B , respectively, as shown in FIG. 6. Samples (transformed coefficients) having a power under the threshold level are omitted as ineffective data (refer to step ST4).

As a result, only samples having a power level over the threshold level are quantized in the circuit 6 and coded by the coding circuit 7, thereby outputting a coded signal (refer to step ST5).

The method can prevent the tone quality from deteriorating by cutting off the signal too much, and the

code from increasing by remaining unnecessary signals, thereby adaptively processing the signals by cutting off the samples having values under the threshold level, whether or not the level of signals are.

Furthermore, as a dynamic range in the processed duration is usually kept at the substantially fixed level, it is possible to reduce the variation of the coding rate in each duration.

[Calculation of power of Input Signal]

The signal power is calculated for a period equal to one block length or more. If the period is lengthened, the coding is late for steep changes of the signal power, so that it is desirable to adopt the method in which the orthogonal transformation length is used as it is. The calculation of the signal power may be performed on the frequency axis after the orthogonal transformation or on the corresponding time axis before the orthogonal transformation.

In former case of the frequency axis, when a point number of the orthogonal transformation denotes as N, the duration power P_i is obtained by an equation,(1) as follows:

$$P_i = \sum_{j=1}^{N} a_{ij}^{2} \quad \text{...} \quad (1)$$

where a_{ij} is represented by a coefficient value corresponding to one of the order j in the duration i.

It is unnecessary to define the threshold level at each frequency. An example in which the threshold level is defined by three straight lines approximating a curve defining the threshold level in the duration from 20 Hz to 20 kHz, is shown in FIG. 7.

In FIG. 7, when frequencies of nodes b and c are respectively set 4 kHz and 14 kHz, obliques of a-b, b-c and c-d are respectively determined as -6.5 db/oct, 5 db/oct and 100 db/oct.

[Determination of the Threshold Level]

The threshold level at 1 kHz is defined as follow:

$(Th)_{1kHz} = 10 \log_{10} P_i - D$ [dB] (2),

thereby setting the entire threshold level corresponding thereto. Where D is determined experimentally.

Next, each power of samples in the respective durations are compared with the threshold level determined above. If the sample power is on or under the threshold level, the power under the threshold is omitted.

[The Feature of this Method]

The feature of the present invention will be explained with reference to FIGS. 2(a), 2(b), 3(a) and 3(b). As is apparent from these figures, it is suitable to select 80 dB as "D" at 1 kHz.

However, a power P_i in a predetermined duration is actually distributed above and below the average level of the figure, and the threshold level rises and falls in each duration unit according thereto. As considered, the influence of samples so omitted is approximated with the condition shown in FIG. 3(a) when the power is under the average level, and with the condition shown in FIG. 3(b) when the power is over the average level. In the former case, the tone quality of the reproduced sound does not deteriorate, and in the latter case, as the total power is sufficiently large in actual, a fine signal is masked so that the quality does not deteriorate.

Furthermore, the code rate decreases because an unnecessary signal samples are omitted in the former case. However, it can be improved by setting a suitable under limit value in the manner that the threshold level does not fall under a predetermined value. In general, there is no great adverse influence on the rise and fall of the threshold level of each duration.

As the result of the estimation of various sound source, the concrete value of D[db] of the equation (2) is set to "D = 80 to 70 [dB]" in an ordinary hearing level, thereby achieving an obbject of this invention,

namely, the quality of reproduced sound does not deteriorate.

Furthermore, the limit value of the threshold level may be set to "$10 \log_{10} a^2$ = -40 -50" at the frequency of 1 kHz.

Claims

1. A method of coding an audio signal by using an orthgonal transformation comprising
a step (ST1, ST2) of calculating the sum total of a power of an input audio signal in a time duration of an orthogonal transform length or over the length;
a step (ST3) of adaptively controlling the minimum audible level as a threshold level corresponding to the power level on a frequency axis after the orthogonal transformation;
a step (ST4) of excepting samples having a power level under the threshold level from orthogonally transformed samples; and
a step (ST5) of quantizing remaining samples.

2. A method of coding an audio signal by using an adaptive orthgonal transformation comprising
a step (ST1, ST2) of calculating the sum total of a power of an input audio signal in a time duration of an orthogonal transform length or over the length;
a step (ST3) of adaptively controlling the minimum audible level as a threshold level corresponding to the power level on a time axis before the orthogonal transformation;
a step (ST4) of excepting samples having a power level under the threshold level from orthogonally transformed samples; and
a step (ST5) of quantizing remaining samples.

3. A coding unit for coding an input audio signal by using an adaptive orthogonal transformation, comprising
window means (1) for dividing said input audio signal into blocks having a predetermined length for an orthhogonal transformation and for setting a minimum audible level from said audio signal;
orthogonal transformation means (2) for transforming said audio signal from said window means (1) by using the orthogonal transformation;
power calculation means (3) for calculating a power level of said audio signal divided by said window means (1);
memory means (4) for storing acoustic characteristic data;
cut-off means (5) for excepting a signal component under said minimum audible level on the basis of said calculated value of said power calculation means (3) and said acoustic characteristics data from said memory means (4);
quantization means (6) for quantizing said signal component output from said cut-off means (5); and
coding means (7) for coding said audio signal after quantizing said audio signal component and for outputting a coded signal to a decoding unit.

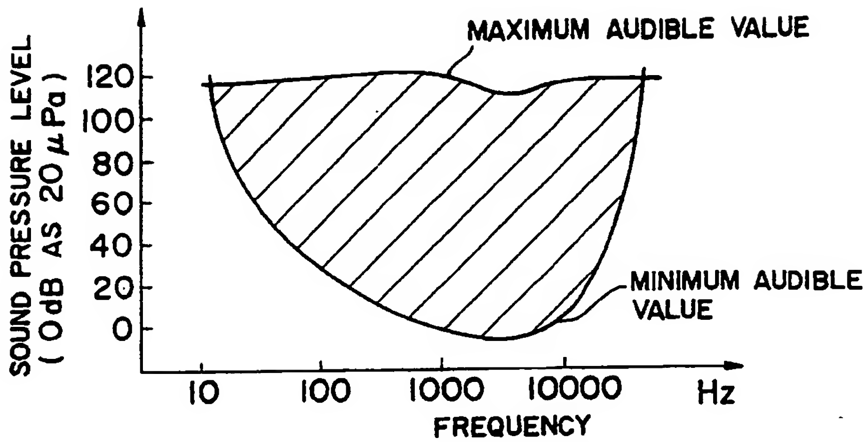


FIG. 1 PRIOR ART

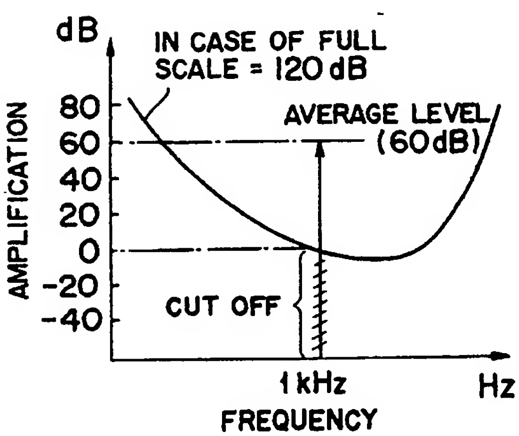


FIG. 2 (a)

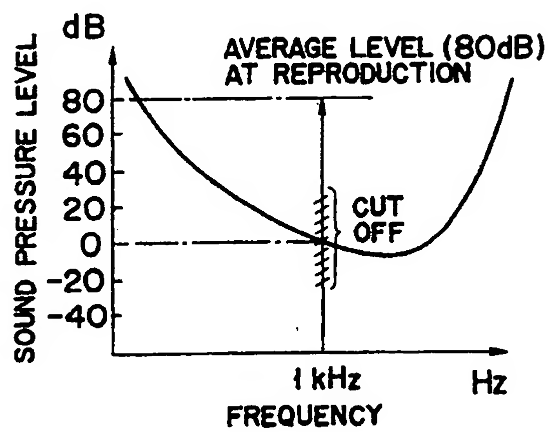


FIG. 2 (b)

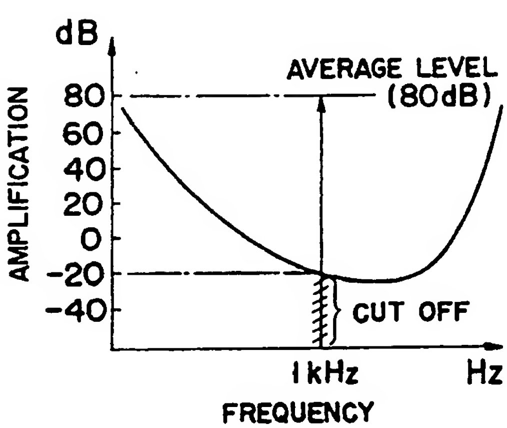


FIG. 3 (a)

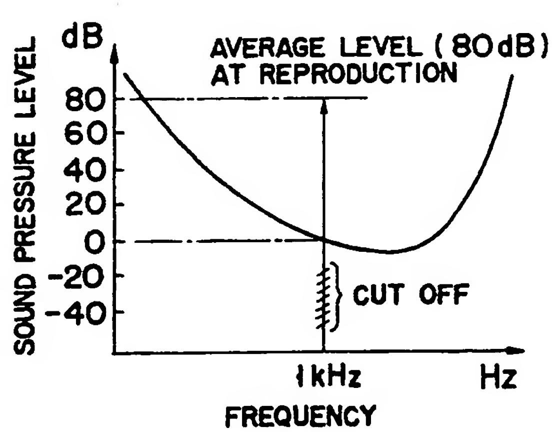


FIG. 3 (b)

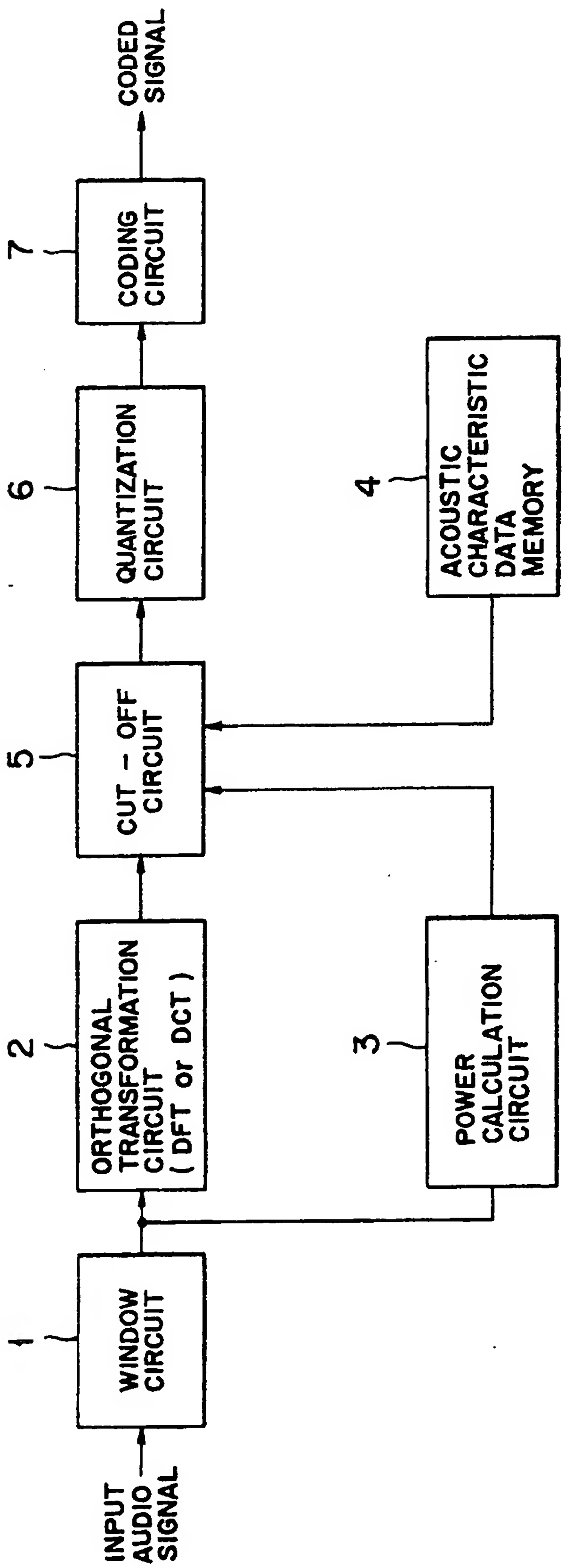

INPUT AUDIO SIGNAL
1
WINDOW CIRCUIT
2
ORTHOGONAL TRANSFORMATION CIRCUIT (DFT or DCT)
3
POWER CALCULATION CIRCUIT
5
CUT – OFF CIRCUIT
4
ACOUSTIC CHARACTERISTIC DATA MEMORY
6
QUANTIZATION CIRCUIT
7
CODING CIRCUIT
CODED SIGNAL

FIG. 4

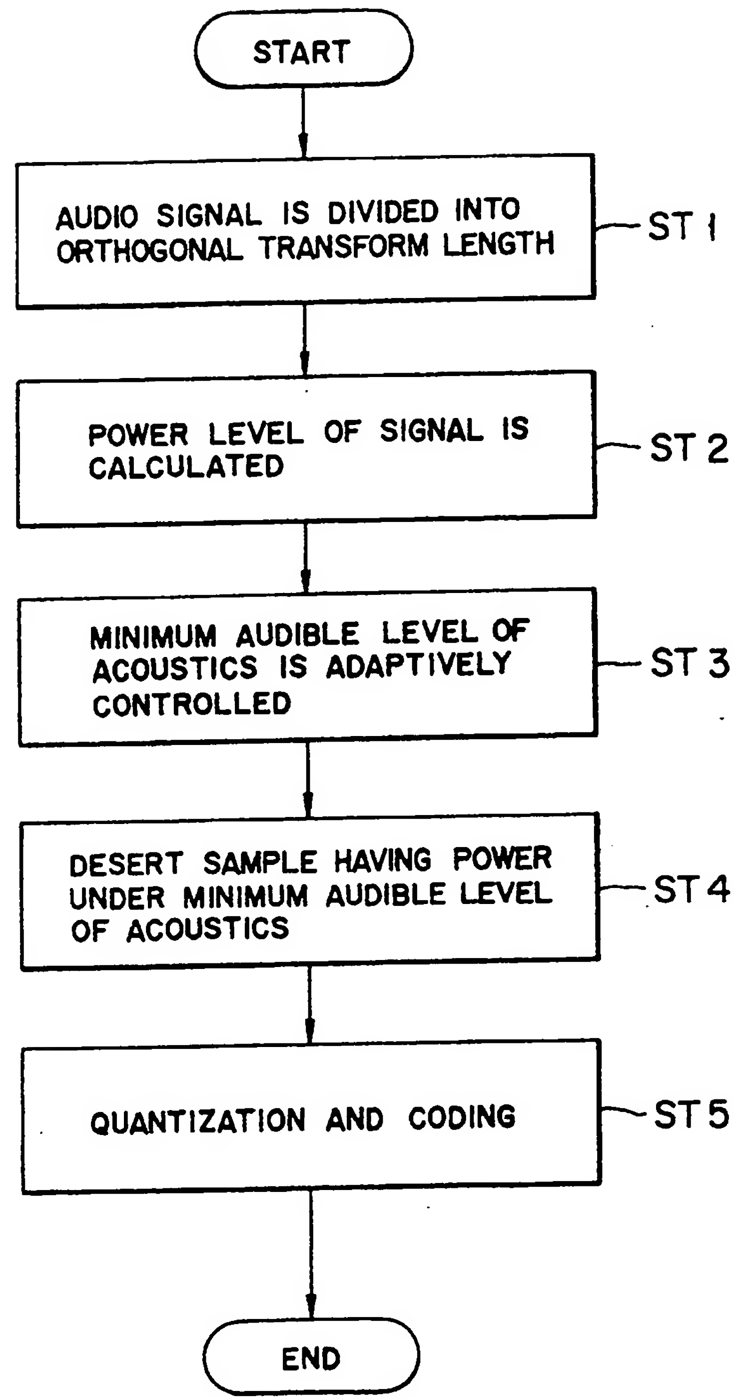
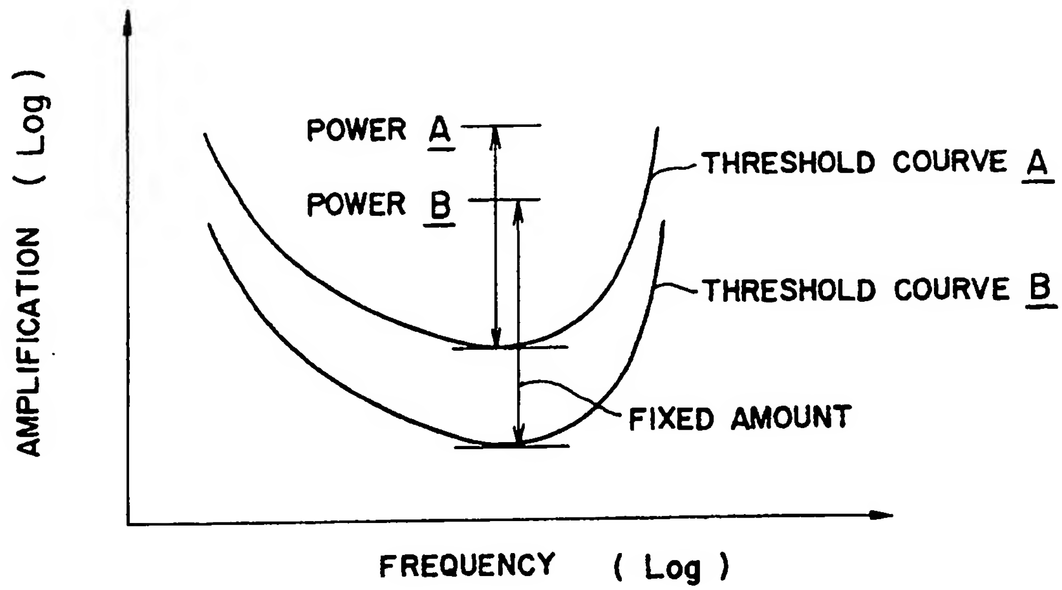
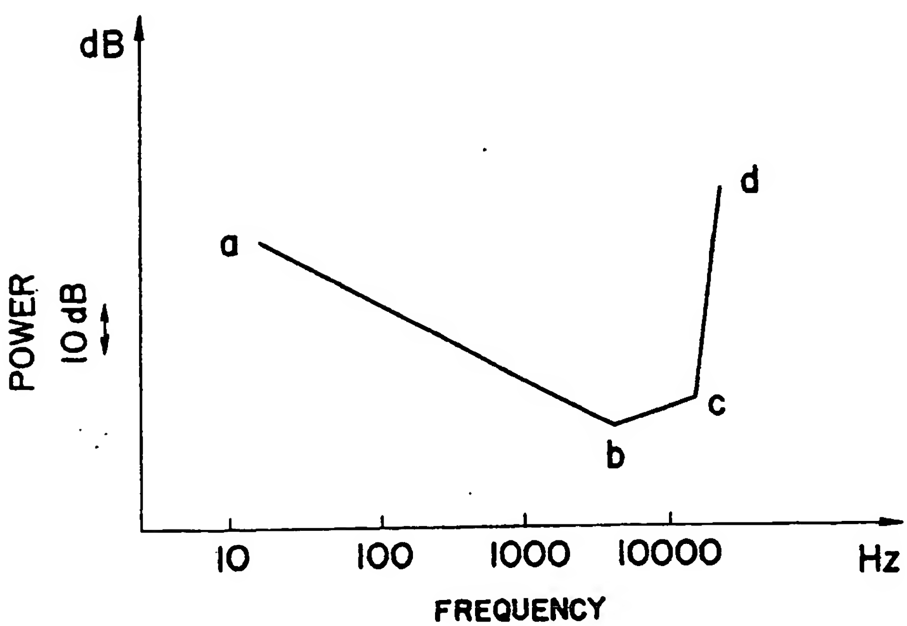

START
AUDIO SIGNAL IS DIVIDED INTO ORTHOGONAL TRANSFORM LENGTH
ST 1
POWER LEVEL OF SIGNAL IS CALCULATED
ST 2
MINIMUM AUDIBLE LEVEL OF ACOUSTICS IS ADAPTIVELY CONTROLLED
ST 3
DESERT SAMPLE HAVING POWER UNDER MINIMUM AUDIBLE LEVEL OF ACOUSTICS
ST 4
QUANTIZATION AND CODING
ST 5
END

FIG. 5


POWER A
POWER B
THRESHOLD COURVE A
THRESHOLD COURVE B
FIXED AMOUNT
AMPLIFICATION (Log)
FREQUENCY (Log)

F I G. 6


dB
POWER
10dB
a
b
c
d
10
100
1000
10000
Hz
FREQUENCY

F I G. 7